

# The inverse of the star-discrepancy problem and the generation of pseudo-random numbers

Josef Dick and Friedrich Pillichshammer

School of Mathematics and Statistics, The University of New South Wales, Sydney  
NSW Australia; email: josef.dick@unsw.edu.au

and

Department of Financial Mathematics, Johannes Kepler University, Linz, Austria;  
email: friedrich.pillichshammer@jku.at

**Abstract.** The inverse of the star-discrepancy problem asks for point sets  $P_{N,s}$  of size  $N$  in the  $s$ -dimensional unit cube  $[0, 1]^s$  whose star-discrepancy  $D^*(P_{N,s})$  satisfies

$$D^*(P_{N,s}) \leq C\sqrt{s/N},$$

where  $C > 0$  is a constant independent of  $N$  and  $s$ . The first existence results in this direction were shown by Heinrich, Novak, Wasilkowski, and Woźniakowski in 2001, and a number of improvements have been shown since then. Until now only proofs that such point sets exist are known. Since such point sets would be useful in applications, the big open problem is to find explicit constructions of suitable point sets  $P_{N,s}$ . We review the current state of the art on this problem and point out some connections to pseudo-random number generators.

## 1 Introduction

The star-discrepancy is a quantitative measure for the irregularity of distribution of a point set  $P_{N,s} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}\}$  in the  $s$ -dimensional unit cube  $[0, 1]^s$ . It is defined as the  $L_\infty$ -norm of the local discrepancy

$$\Delta(\mathbf{t}) := \frac{\#\{n \in \{0, 1, \dots, N-1\} : \mathbf{x}_n \in [\mathbf{0}, \mathbf{t}]\}}{N} - \lambda_s([\mathbf{0}, \mathbf{t}]),$$

for  $\mathbf{t} = (t_1, t_2, \dots, t_s) \in [0, 1]^s$ , where  $[\mathbf{0}, \mathbf{t}] = \prod_{j=1}^s [0, t_j]$  and  $\lambda_s$  denotes the  $s$ -dimensional Lebesgue measure. In other words, the star-discrepancy (or  $L_\infty$ -discrepancy) of  $P_{N,s}$  is

$$D_N^*(P_{N,s}) = \sup_{\mathbf{t} \in [0, 1]^s} |\Delta(\mathbf{t})|.$$

Its significance arises from the classical Koksma-Hlawka inequality [23, 25], which states that

$$\left| \int_{[0, 1]^s} f(\mathbf{x}) \, d\mathbf{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n) \right| \leq V(f) D^*(P_{N,s}), \quad (1)$$

where  $V(f)$  denotes the variation of  $f$  in the sense of Hardy and Krause, see, e.g., [11,28,31]. This is the fundamental error estimate for quasi-Monte Carlo rules  $Q(f) = (1/N) \sum_{n=0}^{N-1} f(x_n)$ .

To provide some insight into this inequality, we prove a simple version of it. Let  $f : [0, 1] \rightarrow \mathbb{R}$  be absolutely continuous, then for  $x \in [0, 1]$  we have

$$f(x) = f(1) - \int_0^1 \mathbf{1}_{[x,1]}(t) f'(t) dt, \quad (2)$$

where  $\mathbf{1}$  denotes the indicator function. Using (2) we obtain

$$\begin{aligned} \int_{[0,1]^s} f(x) dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) &= \int_0^1 f'(t) \left[ \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_{[x_n,1]}(t) - \int_0^1 \mathbf{1}_{[x,1]}(t) dx \right] dt \\ &= \int_0^1 f'(t) \left[ \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_{[0,t]}(x_n) - t \right] dt. \end{aligned}$$

This implies that

$$\left| \int_{[0,1]^s} f(x) dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \leq \int_0^1 |f'(t)| dt \sup_{0 \leq t \leq 1} \left| \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{1}_{[0,t]}(x_n) - t \right|.$$

The right-most expression in the above inequality is simply the star-discrepancy of the point set  $\{x_0, x_1, \dots, x_{N-1}\}$  and for absolutely continuous functions  $f$ , the term  $\int_0^1 |f'(t)| dt$  coincides with the Hardy-Krause variation. This approach can be generalized to the  $s$ -dimensional unit cube  $[0, 1]^s$ , yielding a version of the Koksma-Hlawka inequality (1). To obtain quasi-Monte Carlo rules with small quadrature error, it is therefore of importance to design point sets with small star-discrepancy.

In many papers the star-discrepancy is studied from the view point of its asymptotic behavior in  $N$  (for a fixed dimension  $s$ ). Define the  $N$ th *minimal star-discrepancy* in  $[0, 1]^s$  as

$$\text{disc}(N, s) := \inf_{P_{N,s}} D_N^*(P_{N,s}),$$

where the infimum is extended over all  $N$ -element point sets in  $[0, 1]^s$ . It is well known that  $\text{disc}(N, s)$  behaves like

$$\frac{(\log N)^{(s-1)/2+\delta_s}}{N} \ll_s \text{disc}(N, s) \ll_s \frac{(\log N)^{s-1}}{N}, \quad (3)$$

where  $\delta_s \in (0, 1/2)$  is an unknown quantity depending only on  $s$ . Here  $A \ll_s B$  means that there is a constant  $c_s > 0$  depending only on  $s$  such that  $A \leq c_s B$ . The lower bound was shown by Bilyk, Lacey and Vagharshakyan [5] improving a famous result of Roth [33]. For the upper bound several explicit constructions are known whose star-discrepancy achieves such a bound. See, e.g.,

[11,31]. Thus the upper bound on the  $N$ th minimal star-discrepancy is of order of magnitude  $O(N^{-1+\varepsilon})$  for every  $\varepsilon > 0$ . The problem however is that the function  $N \mapsto (\log N)^{s-1}/N$  does not decrease to zero until  $N > \exp(s-1)$ . For  $N \leq \exp(s-1)$  this function is increasing which means that for  $N$  in this range our discrepancy bound is useless. Even for moderately large dimensions  $s$ , point sets with cardinality  $N > \exp(s-1)$  cannot be used for practical applications in quasi-Monte Carlo rules.

In a number of practical applications one requires point sets with low star-discrepancy whose cardinality  $N$  is not too large compared to  $s$ . This arises naturally, for instance, in estimating the expectation value of a linear functional of the solution to a partial differential equation with random coefficients [29]. From [29, Theorem 8] one can see that in order to reduce the overall error, one needs to balance  $N$  and  $s$  and this balancing is of the form  $N = s^\kappa$ , for some  $0 < \kappa \leq 1$ . Another case arises when the dimension  $s$  is very large. For instance, in some applications from financial mathematics, the dimension  $s$  can be several hundreds, see for instance [32]. If  $s = 100$ , then  $2^s \approx 10^{30}$ . Due to the limitations of the current technology, the number of points  $N$  we can use is much smaller than  $2^s$  in this case. In some instances of these applications, one can still achieve a convergence rate  $N^{-\delta}$  with  $\delta > 1/2$ , but if the problem is more difficult the best we can get is  $N^{-1/2}$ . In these situations one would like to reduce the constant factor which depends on the dimension  $s$ , thus point sets whose star-discrepancy depends only weakly on the dimension  $s$  are useful. At the end of the paper we discuss another situation where the dependence on the dimension is important.

## 2 The inverse of the star-discrepancy problem

We review the current literature on the inverse of the star-discrepancy problem as first studied in [20]. To analyze the problem systematically the so-called *inverse of the star-discrepancy* is defined as

$$N(s, \varepsilon) = \min\{N \in \mathbb{N} : \text{disc}(N, s) \leq \varepsilon\} \quad \text{for } s \in \mathbb{N} \text{ and } \varepsilon \in (0, 1].$$

This is the minimal number of points which is required to achieve a star-discrepancy less than  $\varepsilon$  in dimension  $s$ . The following theorem is the first classic result in this direction.

**Theorem 1 (Heinrich, Novak, Wasilkowski and Woźniakowski [20]).**

*We have*

$$\text{disc}(N, s) \ll \sqrt{\frac{s}{N}} \quad \text{for all } N, s \in \mathbb{N}. \quad (4)$$

*Hence*

$$N(s, \varepsilon) \ll s\varepsilon^{-2} \quad \text{for all } s \in \mathbb{N} \text{ and } \varepsilon > 0.$$

The bound (4) does not achieve the optimal rate of convergence for fixed dimension  $s$  as the number of points  $N$  goes to  $\infty$ . However, the dependence

on the dimension  $s$  is much weaker than in (3). Thus such point sets are more suited for integration problems where the dimension  $s$  is large.

The proof of Theorem 1 is based on the probabilistic method. It is shown that the probability, that the absolute local discrepancy  $|\Delta(\mathbf{t})|$  of a randomly chosen point set is larger than a certain quantity  $\delta$ , is extremely small. Then one applies a union bound over all  $\mathbf{t} \in [0, 1]^s$  and chooses  $\delta$  such that this union bound is strictly less than one, which then implies the result. In this particular instance the authors of [20] used a large deviation inequality for empirical processes on Vapnik-Červonenkis classes due to Talagrand and Haussler. Details can be found in [8, 20]. A simplified proof which leads in addition to explicit constants was given recently by Aistleitner [1].

It is also known that the dependence on the dimension  $s$  of the inverse of the star-discrepancy cannot be improved. Hinrichs [21] proved the existence of constants  $c, \varepsilon_0 > 0$  such that

$$N(s, \varepsilon) \geq c s \varepsilon^{-1} \quad \text{for all } \varepsilon \in (0, \varepsilon_0) \text{ and } s \in \mathbb{N}$$

and  $\text{disc}(N, s) \geq \min(\varepsilon_0, cs/n)$ . The exact dependence of  $N(s, \varepsilon)$  on  $\varepsilon^{-1}$  is still an open question which seems to be very difficult.

Doerr [13] on the other hand showed that the star-discrepancy of a random point set is at least of order  $\sqrt{s/N}$ , which shows that the upper bound of [20] is asymptotically sharp.

A similar but slightly weaker result compared to Theorem 1 is the following:

**Theorem 2 (Heinrich, Novak, Wasilkowski and Woźniakowski [20]).**  
We have

$$\text{disc}(N, s) \ll \sqrt{\frac{s}{N}} \sqrt{\log s + \log N} \quad \text{for all } N, s \in \mathbb{N}, \quad (5)$$

and

$$N(s, \varepsilon) \ll s \varepsilon^{-2} \log(s/\varepsilon) \quad \text{for all } s \in \mathbb{N} \text{ and } \varepsilon > 0.$$

The proof of this result is based on similar ideas as used in the proof of the previous theorem, but instead of the result of Talagrand and Haussler, here the authors of [20] used Hoeffdings inequality, which is an estimate for the deviation from the mean for sums of independent random variables. We give a short sketch of the proof which offers some insights. More details can be found in [11, 20, 30]. *Sketch of the proof of Theorem 2.* Hoeffdings inequality (in the form required here) states that if  $X_0, \dots, X_{N-1}$  are independent real valued random variables with mean zero and  $|X_i| \leq 1$  for  $i = 0, \dots, N-1$  almost surely, then for all  $t > 0$  we have

$$\text{Prob} \left( \left| \sum_{i=0}^{N-1} X_i \right| > t \right) \leq 2 \exp \left( -\frac{t^2}{2N} \right).$$

Now let  $P_{N,s} = \{\mathbf{x}_0, \dots, \mathbf{x}_{N-1}\}$  where  $\mathbf{x}_0, \dots, \mathbf{x}_{N-1}$  are independent and uniformly distributed in  $[0, 1]^s$ . We want to show that

$$\text{Prob}(D_N^*(P_{N,s}) \leq 2\varepsilon) > 0$$

where  $2\varepsilon$  is the right hand side in Theorem 5. That amounts to the task to show that the event

$$|\Delta(\mathbf{x})| > 2\varepsilon \text{ at least for one } \mathbf{x} \in [0, 1]^s$$

has a probability smaller than 1. These are infinitely many constraints, but it can be shown that  $|\Delta(\mathbf{x})| > 2\varepsilon$  implies  $|\Delta(\mathbf{y})| > \varepsilon$  for one of the points in a rectangular equidistant grid of mesh size  $1/m$  with  $m = \lceil s/\varepsilon \rceil$ . Actually, this holds either for the grid point directly below left or up right from  $\mathbf{x}$ . Since this grid has cardinality  $(m+1)^s$ , a union bound shows that it is enough to prove

$$\text{Prob}(|\Delta(\mathbf{x})| > \varepsilon) < (m+1)^{-s}$$

for every  $\mathbf{x} \in [0, 1]^s$ . But now

$$N\Delta(\mathbf{x}) = \sum_{i=0}^{N-1} (\mathbf{1}_{[\mathbf{0}, \mathbf{x})}(\mathbf{x}_i) - \lambda_s([\mathbf{0}, \mathbf{x})))$$

is the sum of the  $N$  random variables  $X_i = \mathbf{1}_{[\mathbf{0}, \mathbf{x})}(\mathbf{x}_i) - \lambda_s([\mathbf{0}, \mathbf{x}))$ , which have mean 0 and obviously satisfy  $|X_i| \leq 1$ . So we can apply Hoeffding's inequality and obtain

$$\text{Prob}(|\Delta(\mathbf{x})| > \varepsilon) = \text{Prob}\left(\left|\sum_{i=0}^{N-1} X_i\right| > N\varepsilon\right) \leq 2 \exp\left(\frac{-N\varepsilon^2}{2}\right) < (m+1)^{-s},$$

where the last inequality is satisfied for the chosen values of the parameters.  $\square$

The results in Theorem 1 and 2 are only existence results. Until now no *explicit* constructions of  $N$ -element point sets  $P_{N,s}$  in  $[0, 1]^s$  for which  $D_N^*(P_{N,s})$  satisfy (4) or (5) are known. A first constructive approach was given by Doerr, Gnewuch and Srivastav [15], which is further improved by Doerr and Gnewuch [14], Doerr, Gnewuch, and Wahlström [17] and Gnewuch, Wahlström and Winzen [18]. There, a deterministic algorithm is presented that constructs point sets  $P_{N,s}$  in  $[0, 1]^s$  satisfying

$$D_N^*(P_{N,s}) \ll \sqrt{\frac{s}{N}} \sqrt{\log(N+1)}$$

in run-time  $O(s \log(sN)(\sigma N)^s)$ , where  $\sigma = \sigma(s) = O((\log s)^2/(s \log \log s)) \rightarrow 0$  as  $s \rightarrow \infty$  and where the implied constants in the  $O$ -notations are independent of  $s$  and  $N$ . However, this is by far too expensive to obtain point sets for high dimensional applications. A slight improvement for the run time is presented in Doerr, Gnewuch, Kritzer and Pillichshammer [16], but this improvement has to be payed with by a worse dependence of the bound on the star-discrepancy on the dimension.

### 3 The weighted star-discrepancy

In the paper [34], Sloan and Woźniakowski introduced the notion of weighted star-discrepancy and proved a “weighted” Koksma-Hlawka inequality. The idea

is that in many applications some projections are more important than others and that this should also be reflected in the quality measure of the point set.

We start with some basic notation: let  $[s] = \{1, 2, \dots, s\}$  denote the set of coordinate indices. Let  $\gamma = (\gamma_j)_{j \geq 1}$  be a sequence of nonnegative reals. For  $\mathbf{u} \subseteq [s]$  we write  $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$ , where the empty product is one by definition. The real number  $\gamma_{\mathbf{u}}$  is the “weight” corresponding to the group of variables given by  $\mathbf{u}$ . Let  $|\mathbf{u}|$  be the cardinality of  $\mathbf{u}$ . For a vector  $\mathbf{z} \in [0, 1]^s$  let  $\mathbf{z}_{\mathbf{u}}$  denote the vector from  $[0, 1]^{|\mathbf{u}|}$  containing the components of  $\mathbf{z}$  whose indices are in  $\mathbf{u}$ . By  $(\mathbf{z}_{\mathbf{u}}, 1)$  we mean the vector  $\mathbf{z}$  from  $[0, 1]^s$  with all components whose indices are not in  $\mathbf{u}$  replaced by 1.

For an  $N$ -element point set  $P_{N,s}$  in  $[0, 1]^s$  and given weights  $\gamma = (\gamma_j)_{j \geq 1}$ , the *weighted star-discrepancy*  $D_{N,\gamma}^*$  is given by

$$D_{N,\gamma}^*(P_{N,s}) = \sup_{\mathbf{z} \in [0,1]^s} \max_{\emptyset \neq \mathbf{u} \subseteq [s]} \gamma_{\mathbf{u}} |\Delta(\mathbf{z}_{\mathbf{u}}, 1)|.$$

If  $\gamma_j = 1$  for all  $j \geq 1$ , then the weighted star-discrepancy coincides with the classical star-discrepancy.

Quite similar to the classical case, we define the  *$N$ th minimal weighted star-discrepancy*

$$\text{disc}_{\gamma}(N, s) = \inf_{P_{N,s}} D_{N,\gamma}^*(P_{N,s})$$

and the *inverse of the weighted star-discrepancy*

$$N_{\gamma}(s, \varepsilon) = \min\{N \in \mathbb{N} : \text{disc}_{\gamma}(N, s) \leq \varepsilon\}.$$

Now we recall two notions of tractability. Tractability means that we control the dependence of the inverse of the weighted star-discrepancy on  $s$  and  $\varepsilon^{-1}$  and rule out the cases for which  $N_{\gamma}(s, \varepsilon)$  depends exponentially on  $s$  or on  $\varepsilon^{-1}$ .

- We say that the weighted star-discrepancy is *polynomially tractable*, if there exist nonnegative real numbers  $\alpha$  and  $\beta$  such that

$$N_{\gamma}(s, \varepsilon) \ll s^{\beta} \varepsilon^{-\alpha} \quad \text{for all } s \in \mathbb{N} \text{ and } \varepsilon \in (0, 1). \quad (6)$$

The infima over all  $\alpha, \beta > 0$  such that (6) holds are called the  $\varepsilon$ -exponent and the  $s$ -exponent, respectively, of polynomial tractability.

- We say that the weighted star-discrepancy is *strongly polynomially tractable*, if there exists a nonnegative real number  $\alpha$  such that

$$N_{\gamma}(s, \varepsilon) \ll \varepsilon^{-\alpha} \quad \text{for all } s \in \mathbb{N} \text{ and } \varepsilon \in (0, 1). \quad (7)$$

The infimum over all  $\alpha > 0$  such that (7) holds is called the  $\varepsilon$ -exponent of strong polynomial tractability.

In both cases the implied constant in the  $\ll$  notation is independent of  $s$  and  $\varepsilon$ .

We collect some known results for the weighted star-discrepancy. The first result is an extension of Theorem 2 to the weighted star-discrepancy.

**Theorem 3 (Hinrichs, Pillichshammer, Schmid [22]).** *We have*

$$\text{disc}_\gamma(N, s) \ll \frac{\sqrt{\log s}}{\sqrt{N}} \max_{\emptyset \neq u \subseteq [s]} \gamma_u \sqrt{|u|}. \quad (8)$$

Note that the result holds for every choice of weights. It is a pure existence result. Under very mild conditions on the weights, Theorem 3 implies polynomial tractability with  $s$ -exponent zero. See [22] for details. A slightly improved and numerically explicit version of Theorem 3 can be found in the recent paper of Aistleitner [2].

Next we state the following result:

**Theorem 4 (Dick, Leobacher, Pillichshammer [9]).** *For every prime number  $p$ , every  $m \in \mathbb{N}$  and for given weights  $\gamma = (\gamma_j)_{j \geq 1}$  with  $\sum_j \gamma_j < \infty$  one can construct (component-by-component) a  $p^m$ -element point set  $P_{p^m, s}$  in  $[0, 1)^s$  such that for every  $\delta > 0$  we have*

$$D_{p^m, \gamma}^*(P_{p^m, s}) \ll_{\gamma, \delta} \frac{1}{p^{m(1-\delta)}}.$$

Note that the point set  $P_{p^m, s}$  from Theorem 4 depends on the choice of weights. The result implies that the weighted star-discrepancy is strongly polynomially tractable with  $\varepsilon$ -exponent equal to one, as long as the weights  $\gamma_j$  are summable. See [9, 10, 11, 22] for more details.

The next result (which follows implicitly from [37]) is about Niederreiter sequences in prime-power base  $q$ . For the definition of Niederreiter sequences we refer to [11, 31].

**Theorem 5 (Wang [37]).** *For the weighted star-discrepancy of the first  $N$  elements  $P_{N, s}$  of an  $s$ -dimensional Niederreiter sequence in prime-power base  $q$  we have*

$$D_{N, \gamma}^*(P_{N, s}) \leq \frac{1}{N} \max_{\emptyset \neq u \subseteq [s]} \prod_{j \in u} [\gamma_j (C j \log(j+q) \log(qN))],$$

with a suitable constant  $C > 0$ .

One can easily deduce from Theorem 5 that the weighted star-discrepancy of the Niederreiter sequence can be bounded independently of the dimension whenever the weights satisfy  $\sum_j \gamma_j j \log j < \infty$ . This implies strong polynomial tractability with  $\varepsilon$ -exponent equal to one. A similar result can be shown for Sobol' sequences and for the Halton sequence (see [36, 37]).

We also have the following recent existence result:

**Theorem 6 (Aistleitner [2]).** *For product weights satisfying  $\sum_j e^{-c\gamma_j^{-2}} < \infty$ , for some  $c > 0$ , we have*

$$\text{disc}_\gamma(N, s) \ll_\gamma \frac{1}{\sqrt{N}} \quad \text{for all } s, N \in \mathbb{N}.$$

Consequently, the weighted star-discrepancy for such weights is strongly polynomially tractable, with  $\varepsilon$ -exponent at most 2.

All results described so far have either been existence results of point sets with small star-discrepancy, or results for point sets with small star-discrepancy which can be obtained via computer search. The Ansatz via computer search remains difficult and is limited to a rather small number of points  $N$  and dimensions  $s$  (in fact, it is known that the computation of the star-discrepancy is  $NP$ -hard as shown by Gnewuch, Srivastav, and Winzen [19], which makes it difficult to obtain good point sets via computer search). To make the random constructions useful in applications, Aistleitner and Hofer [3] show that with probability  $\delta$  one can expect point sets with discrepancy of order  $c(\delta)\sqrt{s/N}$ . Another Ansatz for obtaining explicit constructions is contained in [35].

In the following section we discuss results for explicit constructions of point sets.

#### 4 The weighted star-discrepancy of Korobov's $p$ -sets

Let  $p$  be a prime number. For a nonnegative real number  $x$  let  $\{x\} = x - \lfloor x \rfloor$  denote the fractional part of  $x$ . For vectors we use this operation component-wise.

We consider the so-called  $p$ -sets in  $[0, 1)^s$ , a term which goes back to Hua and Wang [24]:

- Let  $P_{p,s} = \{\mathbf{x}_0, \dots, \mathbf{x}_{p-1}\}$  with

$$\mathbf{x}_n = \left( \left\{ \frac{n}{p} \right\}, \left\{ \frac{n^2}{p} \right\}, \dots, \left\{ \frac{n^s}{p} \right\} \right) \quad \text{for } n = 0, 1, \dots, p-1.$$

The point set  $P_{p,s}$  was introduced by Korobov [27].

- Let  $Q_{p^2,s} = \{\mathbf{x}_0, \dots, \mathbf{x}_{p^2-1}\}$  with

$$\mathbf{x}_n = \left( \left\{ \frac{n}{p^2} \right\}, \left\{ \frac{n^2}{p^2} \right\}, \dots, \left\{ \frac{n^s}{p^2} \right\} \right) \quad \text{for } n = 0, 1, \dots, p^2-1.$$

The point set  $Q_{p,s}$  was introduced by Korobov [26].

- Let  $R_{p^2,s} = \{\mathbf{x}_{a,k} : a, k \in \{0, \dots, p-1\}\}$  with

$$\mathbf{x}_{a,k} = \left( \left\{ \frac{k}{p} \right\}, \left\{ \frac{ak}{p} \right\}, \dots, \left\{ \frac{a^{s-1}k}{p} \right\} \right) \quad \text{for } a, k = 0, 1, \dots, p-1.$$

Note that  $R_{p^2,s}$  is the multi-set union of all Korobov lattice point sets with modulus  $p$ . The point set  $R_{p^2,s}$  was introduced by Hua and Wang (see [24, Section 4.3]).

**Theorem 7 (Dick and Pillichshammer [12]).** *Assume that the weights  $\gamma_j$  are non-increasing.*

1. *If  $\sum_j \gamma_j < \infty$ , then for all  $\delta > 0$  we have*

$$D_{p,\gamma}^*(P_{p,s}) \ll_{\gamma,\delta} \frac{1}{p^{1/2-\delta}}, \quad D_{p^2,\gamma}^*(Q_{p^2,s}) \ll_{\gamma,\delta} \frac{1}{p^{1-\delta}}, \quad \text{and } D_{p^2,\gamma}^*(R_{p^2,s}) \ll_{\gamma,\delta} \frac{1}{p^{1-\delta}},$$



where in all cases the implied constant is independent of  $p$  and  $s$ . This implies strong polynomial tractability.

2. If there exists a real  $\tau > 0$  such that  $\sum_j \gamma_j^\tau < \infty$ , then for all  $\delta > 0$  we have

$$D_{p,\gamma}^*(P_{p,s}) \ll_{\gamma,\delta} \frac{s}{p^{1/2-\delta}}, \quad D_{p^2,\gamma}^*(Q_{p^2,s}) \ll_{\gamma,\delta} \frac{s}{p^{1-\delta}}, \quad \text{and} \quad D_{p^2,\gamma}^*(R_{p^2,s}) \ll_{\gamma,\delta} \frac{s}{p^{1-\delta}},$$

where in all cases the implied constant is independent of  $p$  and  $s$ . This implies polynomial tractability.

The proof of Theorem 7 is based on an Erdős-Turan-Koksma-type inequality for the weighted star-discrepancy and the following estimates for exponential sums. For details we refer to [12].

**Lemma 1.** *Let  $p$  be a prime number and let  $s \in \mathbb{N}$ . Then for all  $h_1, \dots, h_s \in \mathbb{Z}$  such that  $p \nmid h_j$  for at least one  $j \in [s]$  we have*

$$\begin{aligned} \left| \sum_{n=0}^{p-1} \exp(2\pi i(h_1 n + h_2 n^2 + \dots + h_s n^s)/p) \right| &\leq (s-1)\sqrt{p}, \\ \left| \sum_{n=0}^{p^2-1} \exp(2\pi i(h_1 n + h_2 n^2 + \dots + h_s n^s)/p^2) \right| &\leq (s-1)p, \quad \text{and} \\ \left| \sum_{a=0}^{p-1} \sum_{k=0}^{p-1} \exp(2\pi i k(h_1 + h_2 a + \dots + h_s a^{s-1})/p) \right| &\leq (s-1)p. \end{aligned} \tag{9}$$

Inequality (9) is known as Weil bound [38] and is often used in the area of pseudo-random number generation. Constructions related to the  $p$ -sets have also been considered in [7]. All of these constructions are related to the generation of (streams of) pseudo-random numbers (rather than low-discrepancy point sets and sequences). This may not be so surprising since the original argument by Heinrich, et. al. [20] is based on random samples and pseudo-random numbers are designed to mimic randomness. We discuss pseudo-random number generators in the next section more generally.

## 5 Complete uniform distribution and pseudo-random number generators

Pseudo-random number generators are commonly used in computer simulations to replace real random numbers for various reasons. Those point sets are based on deterministic constructions with the aim to mimic randomness. A number of quality criteria are applied to such pseudo-random number generators to assess their quality. One such criterion is complete uniform distribution.

Let  $u_1, u_2, \dots \in [0, 1]$  be a sequence of real numbers. For  $s, N \in \mathbb{N}$  we define

$$\mathbf{u}_n^{(s)} = (u_{(n-1)s+1}, \dots, u_{ns}) \in [0, 1]^s.$$

Then the sequence  $(u_n)_{n \geq 1}$  is completely uniformly distributed if for every  $s \geq 1$

$$\lim_{N \rightarrow \infty} D^*(\{\mathbf{u}_1^{(s)}, \dots, \mathbf{u}_N^{(s)}\}) = 0.$$

The concept of complete uniform distribution measures correlations between successive numbers  $u_i, u_{i+s}, u_{i+2s}, \dots$ . Real random numbers are uncorrelated and thus their discrepancy goes to 0 (in probability), and so one wants pseudo-random numbers with the same property. For instance, the classic construction by van der Corput  $(\phi(n))_{n \geq 0}$  in base 2, given by

$$\phi(n) = \frac{n_0}{2} + \frac{n_1}{2^2} + \dots + \frac{n_m}{2^{m+1}},$$

where  $n$  has dyadic expansion  $n = n_0 + n_1 2 + \dots + n_m 2^m$ , is not completely uniformly distributed, since  $\phi(2n)$  lies in the interval  $[0, 1/2)$ , whereas  $\phi(2n-1)$  lies in the interval  $[1/2, 1)$ .

Markov chain algorithms are a staple tool in statistics and the applied sciences for generating samples from distributions for which only partial information is available. As such they are an important class of algorithms which use pseudo-random number generators. In [6] it was shown that if the random numbers which drive the Markov chain are completely uniformly distributed, then the Markov chain consistently samples the target distribution (i.e. yields the correct result). For instance, [6, Theorem 4] requires pseudo-random numbers  $(u_n)_{n \geq 1}$  such that for every sequence of natural numbers  $(s_N)_{N \geq 1}$  with  $s_N = \mathcal{O}(\log N)$ , we have

$$\lim_{N \rightarrow \infty} D^*(\{\mathbf{u}_1^{(s_N)}, \dots, \mathbf{u}_{s_N}^{(s_N)}\}) = 0.$$

In this case, bounds like (3) are not strong enough due to their dependence on the dimension. Even a bound of the form  $C^s N^{-\delta}$ , with  $C > 1$  and some  $\delta > 0$  which does not depend on the dimension  $s$ , is not strong enough, since for  $s = c \log N$  with  $c > \frac{\log C}{\delta}$ , we have  $C^s N^{-\delta} = C^{c \log N} N^{-\delta} = N^{-\delta + c \log C} \geq 1$  for all  $N \in \mathbb{N}$  and so we do not get any convergence.

Thus it would be interesting for applications to explicitly construct a deterministic sequence  $(u_n)_{n \geq 1}$  such that, say

$$D^*(\{\mathbf{u}_1^{(s)}, \dots, \mathbf{u}_N^{(s)}\}) \leq C \frac{\sqrt{s \log N}}{\sqrt{N}} \quad \text{for all } N, s \in \mathbb{N}.$$

The existence of such a sequence has already been shown in [6, p. 684] and an improvement has been shown in [4]. Such a sequence has good properties when viewed as a pseudo-random sequences but is also useful as a deterministic sequence in quasi-Monte Carlo integration.

## Acknowledgments

J. Dick is support by a Queen Elizabeth II Fellowship from the Australian Research Council. F. Pillichshammer is supported by the Austrian Science Fund (FWF): Project F5509-N26, which is a part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

## References

1. Ch. Aistleitner: Covering numbers, dyadic chaining and discrepancy. *J. Complexity* 27: 531–540, 2011.
2. Ch. Aistleitner: Tractability results for the weighted star-discrepancy. *J. Complexity* 30: 381–391, 2014.
3. Ch. Aistleitner and M. Hofer: Probabilistic discrepancy bound for Monte Carlo point sets. *Math. Comp.* 83: 1373–1381, 2014.
4. Ch. Aistleitner and M. Weimar, Probabilistic star discrepancy bounds for double infinite random matrices. In: J. Dick, F.Y. Kuo, G.W. Peters, and I.H. Sloan, (eds.): *Monte Carlo and quasi-Monte Carlo methods 2012*, Springer, Berlin, 2013.
5. D. Bilyk, M. T. Lacey and A. Vagharshakyan: On the small ball inequality in all dimensions. *J. Funct. Anal.* 254: 2470–2502, 2008.
6. S. Chen, J. Dick, and A.B. Owen, Consistency of Markov chain quasi-Monte Carlo on continuous state spaces. *Ann. Stat.* 39: 673–701, 2011.
7. J. Dick: Numerical integration of Hölder continuous, absolutely convergent Fourier, Fourier cosine, and Walsh series. *J. Approx. Theory* 183: 14–30, 2014.
8. J. Dick, A. Hinrichs, and F. Pillichshammer: Proof techniques in quasi-Monte Carlo theory. Submitted, 2014.
9. J. Dick, G. Leobacher, and F. Pillichshammer: Construction algorithms for digital nets with low weighted star-discrepancy. *SIAM J. Numer. Anal.* 43: 76–95, 2005.
10. J. Dick, H. Niederreiter, and F. Pillichshammer: Weighted star-discrepancy of digital nets in prime bases. In: D. Talay and H. Niederreiter, (eds.): *Monte Carlo and Quasi-Monte Carlo Methods 2004*, Springer, Berlin Heidelberg New York, 2006.
11. J. Dick and F. Pillichshammer: *Digital Nets and Sequences – Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge University Press, Cambridge, 2010.
12. J. Dick and F. Pillichshammer: The weighted star discrepancy of Korobov’s  $p$ -sets. Submitted, 2014.
13. B. Doerr: A lower bound for the discrepancy of a random point set. *J. Complexity* 30: 16–20, 2014.
14. B. Doerr and M. Gnewuch: Construction of low-discrepancy point sets of small size by bracketing covers and dependent randomized rounding. In: *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pp. 299–312. Springer, Berlin, 2007.
15. B. Doerr, M. Gnewuch, and A. Srivastav: Bounds and constructions for the star-discrepancy via  $\delta$ -covers. *J. Complexity* 21: 691–709, 2005.
16. B. Doerr, M. Gnewuch, P. Kritzer, and F. Pillichshammer: Component-by-component construction of low-discrepancy point sets of small size. *Monte Carlo Methods Appl.* 14: 129–149, 2008.
17. B. Doerr, M. Gnewuch, and M. Wahlström: Algorithmic construction of low-discrepancy point sets via dependent randomized rounding. *J. Complexity* 26: 490–507, 2010.
18. M. Gnewuch, M. Wahlström, and C. Winzen: A new randomized algorithm to approximate the star discrepancy based on threshold accepting. *SIAM J. Numer. Anal.* 50: 781–807, 2012.
19. M. Gnewuch, A. Srivastav, C. Winzen: Finding optimal volume subintervals with  $k$ -points and calculating the star discrepancy are NP-hard problems. *J. Complexity* 25: 115–127, 2009.
20. S. Heinrich, E. Novak, G.W. Wasilkowski, and H. Woźniakowski: The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.* 96: 279–302, 2001.

21. A. Hinrichs: Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. *J. Complexity* 20: 477–483, 2004.
22. A. Hinrichs, F. Pillichshammer, and W.Ch. Schmid: Tractability properties of the weighted star-discrepancy. *J. Complexity* 24: 134–143, 2008.
23. E. Hlawka: Funktionen von beschränkter Variation in der Theorie der Gleichverteilung. *Ann. Mat. Pura Appl.* 54: 325–333, 1961.
24. L.K. Hua and Y. Wang: *Applications of Number Theory to Numerical Analysis*. Springer, Berlin, 1981.
25. J.F. Koksma: Een algemeene stelling uit de theorie der gelijkmatige verdeling modulo 1. *Mathematica B (Zutphen)* 11: 7–11, 1942/43.
26. N.M. Korobov: Approximate calculation of repeated integrals by number-theoretical methods. (Russian) *Dokl. Akad. Nauk SSSR (N.S.)* 115: 1062–1065, 1957.
27. N.M. Korobov: *Number-Theoretic Methods in Approximate Analysis*. (Russian) Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1963.
28. L. Kuipers and H. Niederreiter: *Uniform Distribution of Sequences*. John Wiley, New York, 1974; reprint, Dover Publications, Mineola, NY, 2006.
29. F.Y. Kuo, Ch. Schwab, and I.H. Sloan: Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.* 50: 3351–3374, 2012.
30. G. Leobacher and F. Pillichshammer: *Introduction to Quasi-Monte Carlo Integration and Applications*. Compact Textbooks in Mathematics, Birkhäuser, Basel, 2014.
31. H. Niederreiter: *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, 1992.
32. S.H. Paskov and J. Traub: Faster evaluation of financial derivatives. *J. Portfolio Management* 22: 113–120, 1995.
33. K.F. Roth: On irregularities of distribution. *Mathematika* 1: 73–79, 1954.
34. I.H. Sloan and H. Woźniakowski: When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *J. Complexity* 14: 1–33, 1998.
35. V.N. Temlyakov: Greedy-type approximation in Banach spaces and applications. *Constr. Approx.* 21: 257–292, 2005.
36. X. Wang: A constructive approach to strong tractability using quasi-Monte Carlo algorithms. *J. Complexity* 18: 683–701, 2002.
37. X. Wang: Strong tractability of multivariate integration using quasi-Monte Carlo algorithms. *Math. Comp.* 72: 823–838, 2003.
38. A. Weil: On some exponential sums. *Proc. Nat. Acad. Sci. U.S.A.* 34: 204–207, 1948.